Probe

Newsletter for the USDA Plant Genome Research Program

Volume 1, No. 3/4 Fall/Winter 1991

Genome Mapping in Pines Takes Shape

David Neale, Institute of Forest Genetics (IFG), USDA Forest Service, Berkeley, CA
and Ronald Sederoff, North Carolina State, Raleigh, NC

Genome mapping in pine and other forest trees began in the 1970's with the construction of maps based on isozymes. Isozymes are enzyme polymorphisms resulting from amino acid substitutions, which can be detected by gel electrophoresis.

These were the first genetic maps for conifers; large numbers of phenotypic markers do not exist in conifers to construct classical maps. Some of the earliest maps were constructed by Tom Adams at the University of New Hampshire and Oregon State University, Tom Conkle at the USDA Forest Service in Berkeley, and Ray Guries at Yale University and the University of Wisconsin.

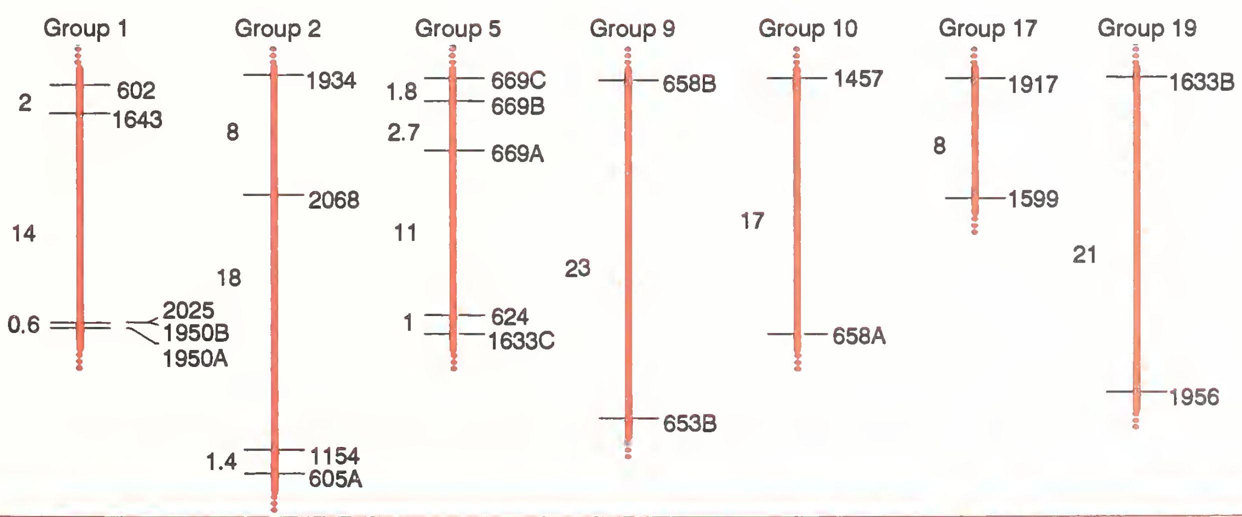
Isozyme linkage maps have now been created for as many as 25 conifer species. These maps were generated rather quickly due to a unique aspect of the genetics of conifer seeds. The nutritive tissue of conifer seeds, the megagametophyte, is haploid and genetically identical to the egg. Thus, linkages between isozyme loci could be established by scoring haploid segregations from individual mother trees, eliminating the need for crosses.

Although the isozyme markers have proven to be extremely useful for population genetic studies, the 30 or so isozyme loci only map a very small segment of the genome. This limitation could easily be overcome with the application of new molecular markers, developed since the 1980's, restriction fragment length polymorphisms (RFLP's), and random amplified polymorphic DNA (RAPD's).

RFLP Map for Loblolly Pine

In 1988, David Neale's lab at the Institute of Forest Genetics (IFG) in Placerville, CA, received funding from ▶

Figure 1. Loblolly Pine RFLP Map



the Forest Biology Program, USDA Competitive Research Grants Office (CRGO), to begin constructing an RFLP map for loblolly pine.

Their approach is to map RFLP's from a single segregating family derived from three generations of outbred matings. The genotypic information on the grandparent trees of the pedigree assists in determining phases in the parent trees. Their map now has 100 RFLP markers located on it, some of which are shown in Figure 1.

In collaboration with Claire Williams and Bob McGraw at the Weyerhaeuser Company, the IFG lab is now mapping on a second 3-generation loblolly pine pedigree with the goal of mapping quantitative trait loci (QTL) coding for wood specific gravity. Specific gravity is a trait of commercial importance and is thought to be under the control of just a small number of genes. The IFG lab is also working with Weyerhaeuser to construct a map for Douglas-fir, another species of commercial importance.

Constructing RAPD Maps

In April 1990, the RAPD mapping technique was introduced by the DuPont Company. RAPD markers are assayed using the polymerase chain reaction (PCR) technique, which is inherently simpler and faster than the Southern blot procedure required for RFLP's.

The drawback of RAPD's is that they are dominant markers as opposed to RFLP's which are codominant. Forest tree mapping researchers quickly realized, however, that the problem of dominance could be overcome by taking advantage of the haploid megagametophytes; RAPD maps could be constructed in the same way as isozyme maps.

Several labs have begun to construct RAPD maps following this approach but Ron Sederoff, David O'Malley, and co-workers at North Carolina State University (NCSU) have made the most progress. In fact, the NCSU lab was able to generate a 191-marker RAPD map for an elite loblolly pine tree in just 2 months (Fig. 2).

Their project was recently funded by the Plant Genome Program of the USDA National Research Initiative Competitive Grants Program (NRICGP). They are planning to develop a theory for mapping QTL's in open-pollinated families from individual trees. This would eliminate the need for making crosses, a very lengthy process in trees.

Third Mapping Effort Funded

The IFG lab has also received funding from the Plant Genome Program to map a major gene for resistance to white pine blister rust. Mike Devey, David Neale, and Bro Kinloch are using the RAPD approach to generate the map from a single mother tree that is heterozygous for the dominant resistance gene. This tree was mated to a homozygous recessive male, therefore, the progeny will segregate 1:1 for resistance and susceptibility. This genetic system will permit mapping the dominant resistance gene based on haploid segregations from the mother tree.

In summary, the three mapping efforts funded by the USDA's CRGO and the NRICGP, in pine are beginning to make significant progress toward mapping the genomes of this very important group of plants. Other pine mapping projects are emerging and will contribute to this effort during the next few years. ◆

TABLE OF CONTENTS

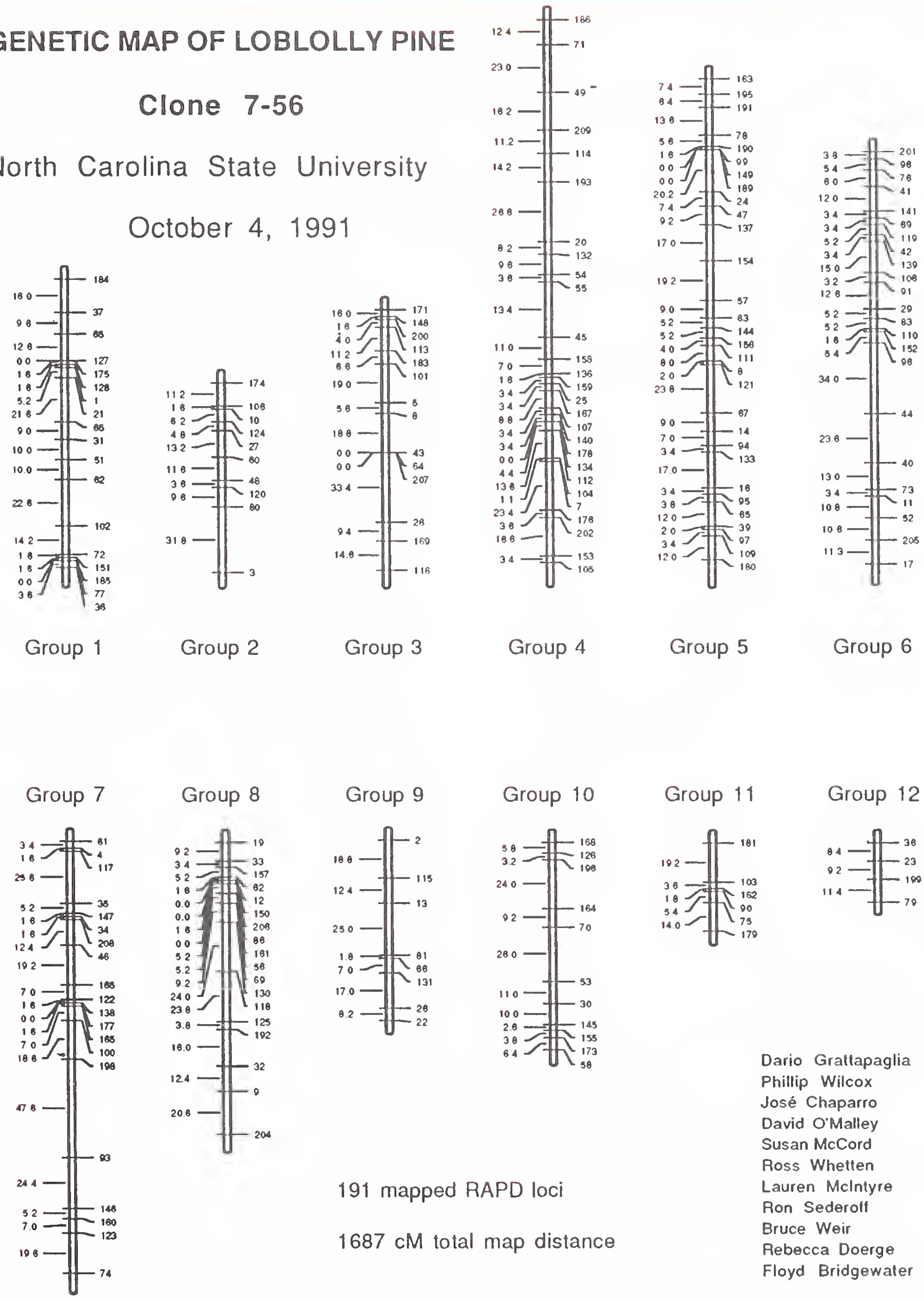
Probe

GENETIC MAP OF LOBLOLLY PINE

Clone 7-56

North Carolina State University

October 4, 1991



191 mapped RAPD loci

1687 cM total map distance

Dario Grattapaglia
Phillip Wilcox
José Chaparro
David O'Malley
Susan McCord
Ross Whetten
Lauren McIntyre
Ron Sederoff
Bruce Weir
Rebecca Doerge
Floyd Bridgewater

Competitive Edge

Increased Funding and New Programs Planned for 1992

Anne Datko, Program Director, National Research Initiative Competitive Grants Program (NRICGP), CSRS, USDA Washington, DC

With approval of the 1992 budget by the President and Congress, funding will again increase for USDA's National Research Initiative Competitive Grants Program (NRICGP). Funds totaling $97.5 million will be available to support the four existing divisions—Natural Resources and the Environment ($18 million); Nutrition, Food Quality, and Health ($6.5 million); Plant Systems ($40 million); and Animal Systems ($25 million)—and to initiate competitive programs in two new divisions—Market, Trade, and Policy ($4 million); and Processes for Adding Value and New Products ($4 million).

1991 Funding and Programs

USDA's former Competitive Research Grants Office (CRGO) accepted responsibility for managing NRICGP when it was established in 1991. At that time, NRICGP received funding at the level of $73 million, a 30-percent increase over 1990's funding for CRGO. Proposal submissions increased 50 percent; 590 awards were made.

The additional applications submitted for established programs plus the large number of applications received for new programs accounted for the increase in 1991. The new programs included Water Quality; Forest, Rangeland, and Crop Ecosystems; Food Safety; Cellular and Molecular Basis of Animal Disease; Animal Molecular Genetics; and Plant Genome.

"Funds totaling $97.5 million will be available to support the four existing divisions... and to initiate competitive programs in two new divisions."

Solicitation for 1992 Proposals

The 1992 Program Description (solicitation for proposals), published in the Federal Register in late November 1991, includes two newly funded Divisions—Markets, Trade, and Policy; and Processes for Adding Value. Also provided is information on the new Agricultural Research Enhancement Award Program. A copy of the Program Description and the Grant Application kit can be obtained from NRICGP, c/o Proposal Services Branch, Cooperative State Research Service, Room 303, Aerospace Center, USDA, Washington, DC 20250-2200. Telephone: (202) 401-5049. Deadlines for all programs are summarized in Table 1.

Peer Reviewers Appreciated

NRICGP solicits proposal reviews from scientists across the country and around the world. Critical and balanced reviews are a crucial component of NRICGP's review process. More than 10,000 reviews were requested in 1991. The effort and time required to prepare thorough and constructive reviews represents a major contribution to NRICGP, the applicant scientists who receive the reviews, and the scientific community as a whole. NRICGP staff extends their appreciation to each scientific peer reviewer who has participated in this process. ◆

Table 1

National Research Initiative Competitive Grants Program

Submission Deadlines

Postmarked Dates	Program Codes	Program Areas	Contacts (202)
January 13, 1992	21.0 23.0 51.1 51.4	Water Quality Forest/Rangeland/Crop Ecosystems Pathology Weed Science	401-6030 401-5114 401-4310 401-4310
January 21, 1992	31.0 54.1 55.0	Human Nutrient Requirements for Optimal Health Photosynthesis and Respiration Alcohol Fuels Research	205-0250 401-6030 401-4310
January 27, 1992	52.1 52.2	Plant Genome Plant Genetic Mechanisms and Molecular Biology	401-4871 401-5042
February 3, 1992	22.1 51.2 51.3	Plant Responses to the Environment Entomology Nematology	401-4871 401-5114 401-5114
February 10, 1992	43.0	Animal Molecular Genetics	401-4399
February 18, 1992	24.0 41.0	Improved Utilization of Wood and Wood Fiber Reproductive Biology of Animals	401-4002 401-6234
February 24, 1992	42.0	Cellular Growth and Developmental Biology of Animals	205-0250
March 9, 1992	53.0 54.2	Plant Growth and Development Nitrogen Fixation/Metabolism	401-5042 401-6030
March 16, 1992	44.0 71.0	Mechanisms of Animal Disease Processing for Value-Added Products	401-4399 401-4002
March 30, 1992	61.0 62.0	Market Assessments, Competitiveness, and Technology Assessments Rural Development	401-4425 401-4425
April 6, 1992	80.1 80.2 80.3	Research Career Enhancement Awards Equipment Grants Seed Grants	401-5114 401-5114 401-5114
April 13, 1992	32.0	Food Safety	401-4399

Nottingham *Arabidopsis* Stock Centre

Dr. Mary Anderson, Director
The Nottingham Arabidopsis *Stock Centre*
School of Biological Sciences
University of Nottingham, United Kingdom

The United Kingdom's Nottingham *Arabidopsis* Stock Centre is part of a multinational, coordinated scientific program to understand the molecular biology, physiology, biochemistry, growth and developmental processes of flowering plants, using *Arabidopsis thaliana* as a model system.

Ideal Research Subject

Although of no economic value, *Arabidopsis* has many attributes that recommend it as a model system. Its potential in this respect is perhaps best exemplified by the plant's nickname, the "Botanical *Drosophila*." So what makes *Arabidopsis* such an ideal subject for research? *Arabidopsis* is a small flowering plant that functions like any other angiosperm. However, it can be grown at high densities—several thousand lines per greenhouse. The plant has a fast cycling time of 6 to 8 weeks so that up to eight generations can be grown in a year, thus facilitating rapid genetic analysis. Using classical genetic analysis, 117 genes have already been mapped on to the five chromosome pairs of the plant.

The plant can also produce many progeny. Following self-pollination, the seed pods, or siliques, each contain more than 30 seeds, one plant can have as many as 200 siliques. Thus, several thousand seeds can be produced from a single plant. The small size of the seeds (about 1 mm long) and plants (30 cm high) simplifies seed mutagenesis and the screening of large populations for new mutations. *Arabidopsis* also has a small genome size, partially because the plant carries a very low level of repetitive deoxyribonucleic acid (DNA). The haploid nuclear genome size is only 100,000 kilobase pairs, which is similar in size to the genome of *Caenorhabditis elegans*, only 7 times larger than that of the yeast *Saccharomyces cerevisiae*, and only 15 times larger than the bacterium *E. coli*! The genome is also significantly smaller than that of any other plant used in research or in agriculture.

Funding and Staff

The Nottingham *Arabidopsis* Stock Centre, located at the University of Nottingham, is funded by the AFRC Plant Molecular Biology Programme; the European Community BRIDGE Programme, "*Arabidopsis* as a tool for isolating genes of agronomic importance;" and the University of Nottingham. The Centre has been funded along with an *Arabidopsis* DNA Stock Centre at Köln, Germany, to serve the European scientific community. The Nottingham Centre is managed by the Head of Centre, Dr. Bernard Mulligan. Dr. Mary Anderson is its newly appointed Director. Two technicians, Mr. Paul Anthony, and Ms. Patricia Fredericks, handle the growth of the plants. The Köln Centre is run by Dr. Jeff Dangl.

Centre's Role

The Nottingham Stock Centre functions as a point of collection, maintenance, cataloging, and distribution of mutant strains and ecotypes of *Arabidopsis*. The intention is that the Centre will work in close association with an American Resource Centre, which is a joint seed and DNA resource centre, based at Ohio State University and run by Dr. Randy Scholl (see his article this issue), to provide an international resource network for the worldwide *Arabidopsis* Program. To fulfill the changing demands of this fast developing area of research, the Centre anticipates working in close association with *Arabidopsis* research-

▶

ers. The Centre has two primary roles—a repository for a diverse collection of *Arabidopsis* lines and a source of information exchange with the scientific community.

The Nottingham *Arabidopsis* Stock Centre already maintains approximately 150 lines of hormone, flowering, biochemical, and form mutants. Many multiple mutants are also available; these are of particular value as marker lines for the mapping of new mutants. Most of these lines have been constructed by Dr. Maarten Koorneef, Wageningen, The Netherlands.

The size and scope of the Centre's holdings will increase dramatically. Over the next year the Centre will incorporate the *Arabidopsis* Information Service (AIS) collection of Professor Kranz, from Frankfurt, which contains over a thousand lines. As the Kranz collection is integrated into the Nottingham collection, a user-friendly database will be created that will carry very detailed information about each line. Initially, the database will not have on-line access, but the Centre's staff will gladly access, upon request, any information required. The database will be compatible with the one established at Ohio State University. The Nottingham Centre intends to produce a short-form Seed List as the Kranz lines are characterized. Eventually, a detailed catalog of all the accessions will be published.

Seed Donations Encouraged

To maximize the genetic diversity held at the Centre, the staff encourages the donation of *Arabidopsis* seeds from as many different sources as possible. If individuals have any novel lines and are prepared for them to be released, the Centre will bulk the seed and offer a safe permanent repository for the line. The Centre is also prepared to hold bulked seed until the details of a particular line have been published. The line will be named and assigned a code for identification purposes. Organizing a central holding center is important in that, at the very least, it will save time and effort in administrating requests and, furthermore, ensure that valuable lines are not lost.

Hopefully, the *Arabidopsis* Stock Centres will be used by all members of the *Arabidopsis* research community.

For more information or to obtain a copy of the Centre's latest Seed List, contact Dr. Mary Anderson, Director, Nottingham *Arabidopsis* Stock Centre, School of Biological Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, U.K. Telephone: +44 602 791216. Fax: +44 602 424270. e-mail: PBZMLH@UK.AC.NOTTINGHAM.CCC.VAX or PBZMLH@VAX.CCC.NOTTINGHAM.AC.UK

◆

Taking Stock

Arabidopsis Biological Resource Center Established at Ohio State

Dr. Randy Scholl
Department of Molecular Genetics
Ohio State University, Columbus

Arabidopsis thaliana is currently taking a place alongside species such as yeast and *Drosophila* as one of the major model research organisms, as discussed in Dr. Mary Anderson's article. As she suggests, it is beneficial to the entire plant and agricultural community for a species so closely related to crop plants to become the focus of such efforts.

More direct applications of this development to crop improvement also should be possible. For example, *Arabidopsis* clones should be useful for both the cloning and alteration of crop genes utilizing biotechnological approaches.

Need for Resource Center

The increase of genetic lines and deoxyribonucleic acid (DNA) clones, associated with the research activity focused on *Arabidopsis*, has become exponential. To preserve these resources and promote their ex- ▶

change, a germplasm and DNA stock facility is being established at Ohio State University (OSU). Funded by the National Science Foundation, The *Arabidopsis* Biological Resource Center at Ohio State (the center), will complement the existing seed stock center at Nottingham, England, and the DNA stock center at Köln, Germany. The creation of *Arabidopsis* stock centers was one of the first goals of the Multinational Coordinated *Arabidopsis thaliana* Genome Research Project.

Responsibilities

OSU's Center will collect and preserve both seed and DNA stocks of *Arabidopsis*, and provide samples upon request. Plans are to establish an on-line database that will contain detailed information about the stocks and species. The staff foresees the database as a user-friendly system that will operate in much the same way as modem microcomputer programs. It will be window-oriented and, when finished next year, should represent a direct and useful connection between the Center and the scientific community.

Since the Center has just been established, the first tasks are to organize its operation, obtain existing large collections, and solicit other stock donations. These operations will be conducted in the next few months. Orders will be taken when the first stocks are ready for distribution, sometime after March 1992.

Meanwhile, OSU staff will be communicating with the *Arabidopsis* community as to the nature of services to be provided, the exact composition of the database, and other aspects of the operating procedures. OSU will use the *Arabidopsis* Bionet electronic news group as well as direct mailings to provide information about the Center. Interested persons may access the electronic news or request that their names be placed on the mailing list.

Staffing for Center

The seed stock collections will be supervised by Dr. Randy Scholl. Dr. Keith Davis of the Ohio State Biotechnology Center will serve as associate director of the Center and supervise the DNA facility. The database will be developed by Dr. Sakti Pramanik from Michigan State University's Computer Science Department.

Plans are to employ four master-level technical assistants at the Center. Their responsibilities will include handling laboratory and greenhouse operations, taking orders, supplying information to patrons, and inputting and retrieving information from the database.

The Center is soliciting applications for these positions from individuals experienced either with *Arabidopsis* genetics or molecular biology techniques. All applications are welcome.

The *Arabidopsis* biological resource centers are certain to be useful to the entire plant research community. The staff at OSU's Center looks forward to working with individuals in the scientific community. For more information, contact OSU at the following address:

The *Arabidopsis* Biological Research Center at Ohio State
1735 Neil Avenue
Columbus, OH 43210
Telephone: 614-292-9371
Fax: 614-292-0603
E-mail: arabidopsis+@osu.edu
Individual Contacts:
R. Scholl - 614-292-1982
K. Davis - 614-292-2115
S. Pramanik - 517-353-3177 ◆

Announcing... New Phone Numbers

Plant Genome Research Program Office

Dr. Jerome Miksche
Program Director
Phone: (301) 504-6029
Fax: (301) 504-6231

Plant Genome Data and Information Center

Dr. Susan McCarthy
Coordinator
Phone: (301) 504-6875
Fax: (301) 504-7098

National Research Initiative Competitive Grants Program

George Jen
Plant Genome
Program Manager
Phone: (202) 401-4264

Touching Base with Ed Coe

Maize Genome Database—Prototype Developing

Dr. Ed Coe, Research Geneticist, Plant Genetics Research Unit
USDA Agricultural Research Service and University of Missouri, Columbia, MO

Corn's significant role in U.S. agricultural productivity is grounded in its innate physiological efficiency, enhanced during this century by breeding for increased productivity, responsiveness to intensive culture, and resistance to pests and stresses.

Less known to the public are its remarkable genetic diversity, the crop's open-book robustness of genetic traits such as kernel variants segregating on an ear of several hundred progeny, and scientists' fascination with its genetic phenomena right from the beginnings of genetic analysis. For example, in addition to his well-known work with peas in the mid-1800's, botanist Gregor Mendel examined segregations in maize. In 1901, German scientist Carl Correns published an extensive monograph on inheritance in maize.

It is little wonder, from this long history of genetic studies, that there are some 670 characterized genes (of which 440 determine "naked eye" polymorphisms; 230 define protein or enzyme polymorphisms). In addition, there are 540 mapped genes; 6,000 mutants awaiting analysis; 1,300 molecular (RFLP) markers in five partly merged maps; 100 cloned sequences of known function; 30 cytologically defined polymorphisms; 880 reciprocal translocations; 84 translocations with the supernumerary B chromosome; and each of 10 trisomics and monosomics.

Maize Database Project

A primary goal of USDA's Plant Genome Research Program is to develop a database system that will contain plant genome information for four agricultural species—soybean, wheat, loblolly pine, and maize.

The intent of the maize database project, currently underway, is to systematize and evaluate maize genetic information, ensure its accuracy, and structure it in a readily accessible database prototype along with data from other crops. A questionnaire sent to maize geneticists last January sought help in defining the essential components and structure of the database; and elicited suggestions, creative ideas, offers of data, and offers of participation in the effort.

Advisory Group Meets

An advisory group of 18 scientists met in St. Louis last April to begin planning the prototype effort. Providing guidance and database expertise for the group were industry-affiliated members (who have had experience with large relational databases), Mary Berlyn (who co-developed the *E. coli* database at Yale University), and Douglas Bigwood (the NAL database manager for USDA's plant genome database project).

The group assigned subcommittees in the areas of Nomenclature and Standards, User Needs and Quality Control, Clone Banks, Quantitative Characters and Descriptors, Germplasm Characterization, Prototype Development, and other high-priority data compilation and collection. (See the shaded box for a list of subcommittees and their respective chairpersons.)

One or more group members have visited database project locations of other species, including Yale University (*E. coli*), the Livermore National Laboratory (human), the Welch Library at Johns Hopkins University (human GDB), the Univer-

▶

sity of Missouri and Washington University (nematodes), Agrigenetics (maize), and the Lawrence Berkeley Laboratory (human).

Project Assignments

Prototype development and implementation are being executed by Letovsky Associates through a research agreement with Stan Letovsky, which includes Yale University's Mary Berlyn as a project cooperator. Initial drafts of the structure, form specifications, and entity relationships have been developed and will soon be provided in revised form.

Priority characterization of germplasm is being carried out for 100 inbred lines across 26 isozymes and 100 RFLP probes by Biogenetic Services, through a research agreement with Alex Kahler. The isozyme analyses have been completed. Reproduction of most of the specific pedigrees has been completed at the University of Missouri. The characterized strains will be deposited in the appropriate ARS germplasm collection.

Scientists Mary Polacco and Marty Sachs, affiliated with the project at Missouri on a half-time basis, are participating in developing the database structure, user needs, quality control, and vocabularies. Mary also is evaluating methods of systematic access and interconnection of genic and mapping data from the literature. Marty is carrying on acquisition and mapping of functionally defined clones, and mapping of reciprocal translocations. The computer specialist on staff at the University of Missouri is geneticist Denis Hancock. ◆

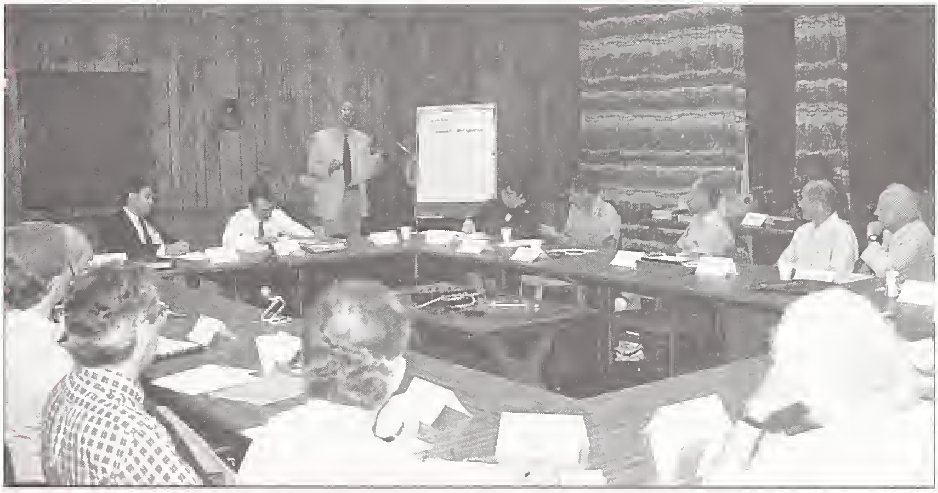
Maize Subcommittees and Chairpersons

Nomenclature and Standards:	Oliver Nelson, University of Wisconsin Phone 608-262-3344, Fax 608-262-2976 nelsonoe@wiscmacc
User Needs and Quality Control:	Ed Coe, USDA-ARS, Univ of Missouri Phone 314-882-2768, Fax 314-875-5359 agrocoe@umcvmb
Clone Banks:	Shiaoman Chao, University of Missouri Phone 314-882-3698, FAX 314-875-5359 and Tim Helentjaris, Univ of Arizona Phone 602-621-8746, Fax 602-621-7186 helnjars@arizrvax
Quantitative Characters and Descriptors:	Jim Coors, University of Wisconsin Phone 608-262-7959, Fax 608-262-5217
Germplasm Characterization:	Charles Stuber, USDA-ARS North Carolina State University Phone 919-737-2289, Fax 919-515-3355
Cytogenetic Data Compilation:	Dave Weber, Illinois State University Phone 309-438-2685, Fax 309-438-3722
Stock Center Cataloging:	Al Kriz, University of Illinois Phone 217-244-6308, Fax 217-333-9817
RFLP Mapping, Pooling of Data:	Tim Helentjaris, University of Arizona (Phone number listed above) and Ben Burr, Brookhaven Nat'l Lab Phone 516-282-3396, Fax 516-282-3407 burr@bnluxl.bnl.gov
Chloroplast:	Steve Rodermel, Iowa State University Phone 515-294-7172
Mitochondrion:	Christiane Fauron, University of Utah Phone 801-581-5192

Home Base

USDA's Plant Genome Database— Collaborative Efforts Continue

Douglas Bigwood, Database Manager
Plant Genome Data and Information Center
National Agricultural Library, USDA
Beltsville, MD



NAL hosted the USDA Plant Genome database technical committee meeting. Discussion was held to plan the scope and implementation for the database.
Photo D. Starr

The Plant Genome Data and Information Center (PGDIC) staff continues to work on the plant genome database design at USDA's National Agricultural Library (NAL) in close collaboration with the other design teams.

Technical Committee Meets

The PGDIC Technical Committee met for the first time July 10th and 11th, 1991. Composed of genetic and information experts, the committee provides technical advice to PGDIC staff who are developing the plant genome database system. As a result of the committee members' input, the database plan of action was improved considerably.

An important point emphasized at the meeting is that individuals involved in developing the database should not attempt to provide all data to all people in all formats. Rather, in the beginning, the project will focus on providing the most crucial data to as many people as possible. In that way, a useful service will be available quickly, then its offerings can be enhanced in the future.

Design Leaps Forward

As reported in the last newsletter, all species groups involved in the database project have undertaken their own database design efforts to meet their specific data processing needs. The wheat, soybean, and pine groups have banded together and contracted with the Lawrence Berkeley Laboratories (LBL) to provide them a database design. LBL has been involved in developing an information system for the Human Genome Project for a number of years. John McCarthy is leading this effort.

The maize group has contracted Stan Letovsky and Mary Berlyn of Yale University's *E. coli* genetic stock center to provide their group with a design. Stan and Mary were responsible for designing and developing the database that supports the center.All of the species groups will be identifying data sources and collecting data to put into their databases. This data eventually will be fed into the database being developed at NAL, where it will be made available to the public. It is critical that all involved in the project closely collaborate so that the data-

continued on page 12 ▶

Other Pursuits

Arabidopsis Genome Research: National Science Foundation Initiative Update

Dr. Machi Dilworth, Program Director
Division of Instrumentation and Resources
National Science Foundation
Washington, DC

The initiative on *Arabidopsis* genome research was proposed in 1989 as part of an initiative on plant sciences by the Biological, Behavioral, and Social Sciences Directorate (BBS) of the National Science Foundation (NSF). Through a series of workshops and meetings that involved scientists active in the international *Arabidopsis* research community, both academic and industrial, and the national funding agencies, a plan for the coordinated *Arabidopsis* genome research project emerged. The plan is described in the 1990 publication "A Long-Range Plan for the Multinational Coordinated *Arabidopsis thaliana* Genome Research Project," (NSF 90-80).

Project Objectives

The primary objective of the project is to encourage a coordinated research effort for the use of *Arabidopsis thaliana* as a model system for studies of the biology of flowering plants. As such, the scope of the project encompasses all areas of plant biology under the purview of the BBS. Specific objectives include (1) identification and characterization of the structure, function, and regulation of genes, (2) development of technologies for plant genome studies, (3) establishment of biological resource

continued on page 13 ▶

Genome Database—cont. from page 11

bases are compatible.So far, collaboration has gone exceedingly well as evidenced by the joint meeting held in Tucson, Arizona, at the Third International Congress of Plant Molecular Biology. At the meeting, the species groups presented their database designs to each other. The designs appear to be similar enough so that areas of incompatibility can be reconciled.

Arabidopsis Joins Effort

A group led by Howard Goodman of Massachusetts General Hospital (MGH) has received funds to develop a database that will include *Arabidopsis* genomic information. Howard and his associates have been working on generating *Arabidopsis* data for many years. Now they will collect and organize this information into a system that will feed the main database here at NAL.

In a separate development, the National Science Foundation (NSF) recently funded a group led by Randy Scholl of Ohio State University (OSU) to develop an *Arabidopsis* stock center. Another group, led by Sakti Pramanik, will develop a database system to support the center. (See Dr. Scholl's article for more information.)

OSU has joined in NAL's database design collaboration. OSU's database will be compatible not only with NAL's database but also with the database at MGH. This means that data submitted to one group will make its way to the other group and eventually to the main database at NAL.

Future Plans

The database design teams will have another joint meeting in mid-January where, hopefully, any remaining incompatibilities in the designs can be reconciled.

In addition, PGDIC staff will begin working on prototypes of applications that will allow end users to access the database. ◆

centers, (4) establishment of an informatics program, (5) development of human resources, and (6) support of workshops and symposia.

Agreement Signed

From the beginning, NSF intended the initiative to be a highly coordinated effort involving international collaboration among *Arabidopsis* researchers, integration of the *Arabidopsis* research community with the general plant biology research community, and close collaboration with other genome research projects. To facilitate this coordination, an interagency agreement was signed in June 1990 by the National Institutes of Health, the Department of Energy, the U.S. Department of Agriculture, and NSF to collaborate and coordinate the *Arabidopsis* genome research efforts in the United States. The agreement designated NSF as the lead agency.

Proposal Reviews

In FY 1991, BBS received an additional funding of $4.4 million to support the *Arabidopsis* initiative. Since the existing BBS program structures were able to adequately review all the expected proposals that fell within the scope of the *Arabidopsis* genome research, no special review panels or programs were created for the new initiative. Each proposal was submitted to and reviewed by the appropriate programs.

Requests for supplements to the existing NSF grants for additional graduate or postdoctoral students or short-term research visits were reviewed by the program who had responsibility for the existing grant.

Infrastructural support such as the resource centers, informatics needs, postdoctoral fellowships, and the development of research tools and resources were handled by the Division of Instrumentation and Resources. The use of existing programs for review ensured that only those proposals that competed well with the rest of the proposals (i.e., non-*Arabidopsis* proposals which are the majority) would receive support under the new *Arabidopsis* initiative.

Support Continues

As various BBS programs had been receiving an increasing number of proposals that would fall under the purview of the *Arabidopsis* initiative, it was expected, and later shown to be true, that even with additional funds, all worthy proposals in the *Arabidopsis* genome research area could not be supported in FY 1991. NSF will continue the support of the *Arabidopsis* genome research initiative in FY 1992.

In addition to support for various scientific activities, NSF has supported the *Arabidopsis* genome research project by providing funds for meetings of the Multinational Science Steering Committee and for publications and distribution of the annual progress report. Through the Steering Committee, various national funding agencies have been involved in support of this international effort. As the designated lead agency for the U.S. effort, NSF maintains contact with its counterparts in other countries in order to facilitate the coordination efforts by the scientists.

Progress Documented

The first annual progress report for the Multinational Coordinated *Arabidopsis thaliana* Genome Research Project was published in June 1991 (NSF 91-60). Remarkable progress is documented in the report. The progress is continuing at an accelerated pace. Most likely, the next annual progress report will have many new advances to report, particularly in the area of informatics.

The rapid advances of the *Arabidopsis* genome project can be attributed, in part, to the willingness of scientists to communicate and share information and resources. NSF staff hopes the *Arabidopsis* Genome Research Project will not only contribute to the advances in plant biology but also will serve as a model for multinational collaboration in plant biology research. ◆

Send your publication requests to.....

National Science Foundation
Publications Office
1800 G St., N.W.
Washington, DC 20550
Phone: (202) 357-7861

Touching Base with Randy Shoemaker

Soybean Genome Database Project Enters New Phase of Development

Dr. Randy C. Shoemaker
Research Geneticist, Field Crops Research Unit
Iowa State University
Agricultural Research Service, USDA

The Soybean Genome Database (GDB) Project is entering a new phase of development. After months of discussion and planning by the soybean team, the early stages of initiating a working prototype database are now underway. A version of the model should be ready for demonstration by late March 1992.

Interagency Agreement

Participants in the Soybean GDB Project recently entered into an interagency agreement with the Lawrence Berkeley Laboratory's (LBL's) Genome Computing Group. With John McCarthy and Suzanna Lewis as principle investigators, LBL's Information and Computing Sciences Division will work with soybean research scientists in developing a prototype soybean genome database.

LBL staff will use database design tools to define entities, relationships, and attributes for the soybean genome database. They will then use companion software tools to translate those high-level definitions into a relational database definition suitable for Sybase (a commercial software product for relational databases).

Soybean GDB Project experts, Wheat GDB Project experts, and LBL staff will work together in an attempt to identify common ground between the needs of the Soybean GDB Project and the needs of the Wheat GDB Project. Every effort will be made to create a "fused" database design that will meet the requirements of a wide variety of plants.

Working Meeting Held

At a recent meeting held at Iowa State University, participants focused on priority targets to be included in the initial database prototype. In addition to focusing on soybean germplasm, genetic diversity, and genetic maps, the group decided the database should include information on soybean pathogens and pathology. Roger Boerma, University of Georgia, was asked to take the lead in determining the priorities for the pathology and pathogen database.

In attendance at the meeting were LBL staff member (Suzanna Lewis), two representatives from the Wheat GDB Project (Bob Graybosch and Olin Anderson), a plant genome database advisor (Mary Berlyn), two representatives from the National Agricultural Library's Plant Genome Data and Information Center (Douglas Bigwood and Rose Broome), the assistant curator of the soybean germplasm collection, soybean scientists, and computer consultants.

Efforts Continue

In early September, several soybean researchers visited LBL staff to confer with them on the organization of the database. This meeting has been followed up with assignments of specific attributes to each of the topics included in the initial prototype. These ideas and organizational formats were presented to the Plant GDB Project leaders at the Plant Genome Database Design Conference held in conjunction with the Third International Congress of Plant Molecular Biology, in Tucson, Arizona. ◆

International Resources on the Release of Organisms Into the Environment

Dr. Mark Segal, Member, International Resources on the Release of Organisms (IRRO) Steering Committee
Environmental Protection Agency
Washington, DC

The release of organisms into the environment is a topic of considerable interest to scientists, public interest groups, industry, and government and nongovernmental agencies alike. There is a widely felt need for reliable, comprehensive information on all aspects of this topic. Although several individual efforts have been initiated, a unified global information resource does not exist.

Collaborative Effort Initiated

In 1990, the United Nations Environmental Programme (UNEP) recognized the need for such a resource. UNEP invited the Microbial Strain Data Network (MSDN) to organize workshops to examine the needs and specifications for a worldwide information system with the possible aim of establishing such a system. UNEP provided seed funding to initiate the process. MSDN solicited participation from other organizations.

Additional funds for the workshops came from the U.S. Environmental Protection Agency (OTS/HERD), the U.S. Department of Agriculture (APHIS/BBEP), Environment Canada (Commercial Chemicals Branch), and the Commission of the European Communities (DGX11/F/1).

Because of travel restrictions, the first workshop, held March 11-15, 1991, consisted of simultaneous sessions at two locations—Vienna, Austria, and Rockville, Maryland, USA. The workshop brought together experts in microbiology, industrial experiences, release regulatory matters, database and network development, and management.

The meetings were linked by electronic mail, prerecorded videotapes, and teleconferences, and by an on-going computer conference where those unable to attend the workshop sites could contribute. These sessions firmly established the need for, and broadly outlined an approach to, building the system.

Second Workshop Held

Nottingham, United Kingdom, was the site for a second workshop held August 28-29, 1991, in conjunction with the Release of Genetically Engineered Microorganisms (REGEM-II) meeting. It brought together participants from both previous sessions and confirmed the network approach recommended by the earlier groups. The Nottingham workshop also set the stage for implementing the system by drawing the following conclusions:

- An integrated information resource on the introduction of organisms into the environment should be established.
- The resource should be international in scope and accessibility.
- Information on both nonmodified and genetically modified organisms should be included.
- The resource should be not-for-profit.

▶

- Data should be made freely available as far as possible; confidential commercial information would be the major exclusion.
- The resource should be a distributed network linking existing resources.
- New databases should be developed only to fill gaps in information, thereby avoiding duplication.
- The separate information resources should be linked by gateways where appropriate.
- Usage should be facilitated by the development of common interfaces.
- The resource should be flexible to accommodate different information sources already in operation (that may have existing distribution policies) and to adjust to new circumstances that may arise.
- Regional help desks and training should be important parts of the resource.

Consensus of Participants

The strong consensus of workshop participants was that the success of the resource depended on the close collaboration between the scientific, regulated, and regulatory communities. The scope of the resource was seen to cut across traditional disciplinary lines and potentially to support such diverse interests as biotechnology, biodiversity, and bioremediation.

Moreover, the participants agreed that the initiative was relevant to various international environmental efforts and that the recommendations of the workshop should be conveyed to the 1992 United Nations Conference on the Environment and Development (UNCED) planned for Rio de Janeiro and to other related programs under consideration.

These programs include initiatives such as those of various components of the International Council of Scientific Unions (e.g., IUBS, IUMS, WFCC), the Organization for Economic Cooperation and Development (OECD), the Commission of the European Community, and various U.S. agencies.

Steering Committee Formed

An international Steering Committee for IRRO (International Resources on the Release of Organisms into the Environment) has been formed to guide the development of the data resource and provide scientific and technical oversight. MSDN provides networking expertise and administrative support. The Steering Committee anticipates that a demonstration resource, consisting of varied information sources and services, will be available to the interested public in time for the UNCED meeting.

The steering committee members come from four continents and from island nations. They represent various scientific disciplines from academic, industrial, and governmental backgrounds. This group by design include users of the resource, regulatory professionals, classical biologists, private sector consultants, computer scientists, ecologists, and public health administrators. Supplementing this group will be volunteers who will be organized into special purpose working groups.

Electronic Network

IRRO is not envisioned as a single database, *per se*. It will be an electronic network that will provide centralized access to existing data sources in different regions of the world. The network will use modern telecommunications systems and set up gateways and interfaces. It will provide a single contact point for all those studying the release of organisms into the environment.

The developers of the new network see IRRO as having a broadly based user community. Each component of the community will probably focus on a subset of the resources made available through, or among, disciplines due to the considerable overlap in data resource interests among the varied user groups. User groups will likely include the following:

1. Developers of releases: commercial, academic or governmental.
2. Ecologists and others studying release phenomena.
3. Funding agencies.
4. Regulatory agencies.
5. Public interest groups.
6. Scientific, technical, and commercial associations or organizations.
7. Systematics collections of biotic materials.

Resource Scope

In keeping with this diverse user community, the scope of the resource will be similarly wide ranging. The network will include access to information on the kinds of released organisms and their characteristics, and those of related organisms. Many varieties of organisms currently are being released or consid-

IRRO CONTACTS

Inquiries can be made to the following partial list of members of the Steering Committee and/or the Secretariat:

IRRO/MSDN Secretariat
307 Huntingdon Road
Cambridge CB3 OJX, UK
Tel: 44 223 277628
E-mail BT Gold 75:DBI0005

Ms. Cindy Callahan
USDA, APHIS, BBEP
Federal Building, Room 854A
Hyattsville, MD 20782, USA
Tel: 1 301 436 5140
Fax: 1 301 436 8669
E-mail BT North America
42:CDT0602

Dr. Vanderlei Canhos
Base de Dados Tropical
R. Latino Coelho, 1301
13085 Campinas SP
Brazil
Tel: 55 192 42 7827 / 7022 ext 72
Fax: 55 192 42 7827
E-mail BT North America
142:CDT0094
burc.bitnet
ccvax.unicamp.ansp.br

Dr. Peter Kearns
OECD
2 rue Andre-Pascal
75775 Paris Cedex 16
France
Tel: 33 1 4524 9975
Fax: 33 1 4524 1675

Dr. Brain Kirsop
Biostrategy Associates
10 Waterside
Ely, CB7 4AZ, UK
Tel: 44 353 663562
Fax: 44 353 663436
E-mail Telecom Gold
75:DBI0219

Prof. Ashok Kolaskar
Bioinformatics DIC
University of Poona,
Pune 411 007, India
Tel: 91 212 335039
Fax: 91 212 330087
Telex: 145629 GENE IN
E-mail Telecom Gold
75:DBI0295

Dr. Anne-Marie Prieels
Av de l'Observatoire 2
B-1180 Brussels
Belguim
Tel: 32 2 374 0076
Fax: 32 2 374 0076

Dr. Mark Segal
Environmental Protection Agency
EPQ-QTS (TS796)
401 Street SW
Washington, DC 20460, USA
Tel: 1 202 382 3389
E-mail BT North America
142:CDT0158

Dr. Naresh Singh
Caribbean Environmental Health Institute
P. O. Box 1111, Morne Fortune, Castries
St. Lucia
Tel: 1 809 452 2501
Fax: 1 809 45 32721

rently are being released or considered for that purpose. The system will not attempt to discriminate among data sources based on the category of organism, but rather, will attempt to provide access to data on microbes and macrobiota, genetically engineered or not. It is designed to provide information to aid a developer choosing a candidate organism from a list provided by a participating source of those organisms, an investigator evaluating an organism related to one intended for release, or an ecologist predicting the progress and likelihood for success of the release.

Additionally, IRRO will help users locate information on releases planned or already in progress. Descriptions of releases including site characteristics, methods of site selection, approaches to monitoring, and observations of effects may be listed in reports that could be located through IRRO. Information on the people involved in releases could be captured and made available to those who need to consult on those events.

Activities of government, scientific, or industrial bodies that effect the progress of release-related activities could be listed and cross referenced for the benefit of IRRO users. Expert lists, schedules of key meetings and conferences, and sources of recent publications—all may be made available via electronic mail, bulletin boards, or computer conferences. These non-database functions should dramatically improve communications among disparate groups concerned with organism releases. ◆

HyperGene: Software for DNA-Based "Graphical Genotypes"

Nevin Dale Young, Assist. Professor
Department of Plant Pathology
University of Minnesota
St. Paul, MO

As scientists develop new techniques to examine the genetic constitution of individual plants, there is a growing need for simple methods to display the information. For example, detailed genetic maps based on deoxyribonucleic acid (DNA) markers (such as RFLP's) make it possible to describe the ancestry of an individual throughout its entire genome. While it is widely agreed that this kind of information will make it possible to develop better plants faster, scientists face a serious problem of effectively making sense of so much information. One computer program that addresses this problem is "HyperGene."

Graphical Genotype Display

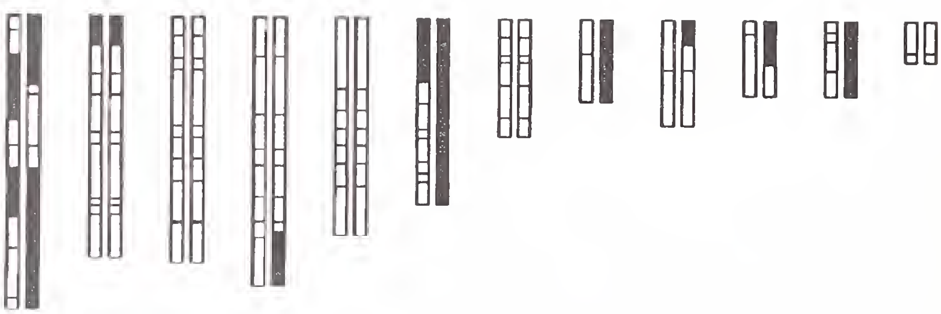
HyperGene is a Macintosh program that uses DNA marker data to create a graphical image for the genomic constitution of an individual. The program actually draws a genetic map showing segments of DNA from different ancestors in different shades or colors. This way of looking at a mosaic of inherited chromosomal regions is called a "graphical genotype."

HyperGene was originally developed by Dr. Nevin Young working with Steve Tanksley at Cornell University. Dr. Young, who has since moved to the University of Minnesota, is now expanding the program to have more powerful display routines and improved decision-making capacities. The program is also being adapted so that is will work in conjunction with USDA's Plant Genome Database Project. As a model system for developing these new features, Dr. Young and his associates are focusing on soybean and the soybean genome mapping database.

Selecting Target Individuals

In addition to displaying graphical genotypes, HyperGene also makes it possible to select target individuals out of a population by simply "painting" the desired genome using a computer mouse. If more than one individual in the population has the desired configuration, HyperGene can compare the selected individuals and pinpoint the best for future breeding.

HyperGene, and programs like it, provide a powerful tool for applying genetic data to crop improvement. The program considers the genome as a whole by integrating the results from many discrete DNA markers. In this way, scientists can get the "big picture" instead of being burdened by an overload of detail. And since users can precisely describe a target genome in terms of DNA genotype, HyperGene should also make plant breeding a more deterministic process. (See box on page 19 for disk information.) ◆



A Graphical Genotype for an F2 individual. Dark regions indicate DNA segments derived from one parent; light regions DNA regions from the other parent.

USDA Grants Boost Plant Genome Research

John D. Copeland
National Center for Agricultural Law Research and Information
University of Arkansas
Fayetteville, Arkansas

As a result of selective plant breeding, worldwide crop production dramatically increased during the past half century. But the recent emergence of a major technological breakthrough in plant genetics promises to revolutionize the development of plant varieties.

Plant gene or *genome mapping* permits scientists to identify the actual plant genes that control physiological growth, development, and productivity in virtually any plant species.

By mapping the gene codes of certain plants, such as rice, wheat, and corn, scientists can isolate traits that protect plants from diseases and pests, as well as those traits that produce higher yields. Eventually, scientists hope to cut in half the time it takes to get desired traits in plants.

Grant Program Established

Because of plant genome mapping's potential impact on agriculture, horticulture, and the world food system, American scientists have called on the U.S. Congress to establish programs to facilitate plant genome mapping. In response, the U.S. Department of Agriculture (USDA) has launched a plant genome mapping program. The 1991 farm bill establishes a research grant program to support basic and applied research and technology in the development of plant genome structures and functions. 7 U.S. Code Annotated (U.S.C.A.) § 450 (i) (1991). The program's goal is to make the United States a leader in biotechnology and to develop profitable crop varieties while creating a positive effect on the environment.

The program authorizes the Secretary of Agriculture to award competitive grants for periods up to 5 years for research projects in plant genomic mapping. Grant recipients can be State agricultural experiment stations, colleges and universities, other research institutions and organizations, Federal agencies, and even private organizations, corporations, or individuals. *Id.* § 450i(b)(1).

High-Priority Research

To the greatest extent possible the grants are allocated to high-priority research, taking into consideration the determination made by the Joint Council on Food and Agricultural Sciences and the National Agricultural Research and Extension Users Advisory Board. *Id.* § 450i(b)(1)

"High-priority research" is defined as basic and applied research that focuses on national and regional research needs in plant systems, including plant genome structure and function, molecular and cellular genetics and plant biotechnology, plant-pest interactions and biocontrol systems, crop plant response to environmental stresses, improved nutrient qualities of plant products, and new food and industrial uses of plant products. High-priority research also includes the methods to transfer such research to on-farm or in-market practice. *Id.* § 450i(b)(2)(A).

Restrictions

There are a number of critical restrictions on the use of high-priority research grant funds. For example, such funds cannot be used for planning, repair, rehabilitation, acquisition, or construction of a building or facility. *Id.* § 450(i)(7). Also, where appropriate, the grants ▶

must be consistent with the development of sustainable agriculture. *Id.* § 450(i)(j).

The Secretary is authorized to formulate rules and regulations to govern appropriate licensing and patent arrangements, copyright fees, royalties, or other fee arrangements from the sales of products and their uses, applications, technology, or other processes developed through the plant genome mapping program.

Funds Authorized

The Secretary of Agriculture is required to file a written report with Congress by January 1 of each year describing the policies, priorities, and operations of the grant program. *Id.* § 450(i)(b)(9). Congress authorized $150 million for the program for 1991 and an additional $275 million for 1992. Funds have been authorized through 1995 when the amount is to reach $500 million. *Id.* § 450(i)(b)(10). ◆

Send your formatted disk for Hypergene

Hypergene is written for the Macintosh computer. The program requires at least 2 megabytes of RAM and a math coprocessor. All non-profit scientists may receive a copy at no cost by sending a formatted Macintosh diskette and a stamped, self-addressed diskette mailer to:

Nevin Dale Young
Department of Plant Pathology
495 Borlaug Hall
University of Minnesota
St. Paul, MN 55108

Off the Wire

Bibliography Available on Computational Molecular Biology

Dr. Sarah Barron
Group Leader, GenTools™ Project
University of Texas System Center for High Performance Computing
Austin, TX

The field of computational molecular biology and genetics is expanding at an enormous rate. Journals such as **CABIOS and Nucleic Acids Research** now routinely publish articles on the computational and mathematical aspects of biology.

A database of bibliographic references on computational algorithms in molecular biology and genetics is maintained by staff of the GenTools™ Project, University of Texas System Center For High Performance Computing.

The GenTools™ **CMB-Bibliography** effort focuses upon computer and mathematical aspects of molecular biology and genetics (interpreted in a broad and liberal sense). Authors are solicited for their additions / corrections to this ongoing effort. The bibliography currently contains approximately 2,500 citations, but continues to expand on a daily basis.

The **CMB-Bibliography** is available to the general public. To obtain the bibliography contact:

Sarah Barron
GenTools™ Project Leader
UT System Center for High Performance Computing
Balcones Research Center
1.154CMS
10100 Burnet Road
Austin, TX 87812

or send an e-mail request to:
GENTOOLS@CHPC.UTEXAS.EDU

Contributions of citations, reprints, preprints, and suggestions, as well as requests for further information about the GenTools™ Project, may be sent to Dr. Sarah Barron at the above address. ◆

The National Research Initiative Competitive Grants Program has a new Plant Genome Program Manager.

George Jen will step in for Anne Datko on January 27, 1992. He can be reached at *(202) 401-4264*. Check out the next issue of *Probe* for more details.

Cloning and Amplifying Large Genomic DNA Fragments

Dr. Phillip R. Buzby
E. I. Du Pont Biotechnology Systems R&D
N. Billerica, MA

The construction of detailed gene-linkage maps of complex genomes facilitates the localization of genes to specific regions on chromosomes. Further, precise loclalization of specific genes requires cloning of high-molecular-weight DNA if it is to be done efficiently. Two methods commonly used are to either insert large genomic DNA fragments into cosmid vectors, which are then packaged into Phage λ heads and propagated in *Escherichia coli*,[1] or into yeast artificial chromosome (YAC) vectors propagated in *Saccharomyces cerevisiae*.[2] The use of cosmids results in high transformation efficiencies, but insert size is limited by the size of the Phage λ head to 45-48 kilibase pairs (kbp). The advantage of using YAC vectors is their ability to accommodate DNA as large as 200-800 kbp. Disadvantages are the low transformation efficiency, which decreases with insert size; the need to process transformants individually prior to screening; and the difficulty in obtaining large amounts of recombinant DNA from transformed cells.[3]

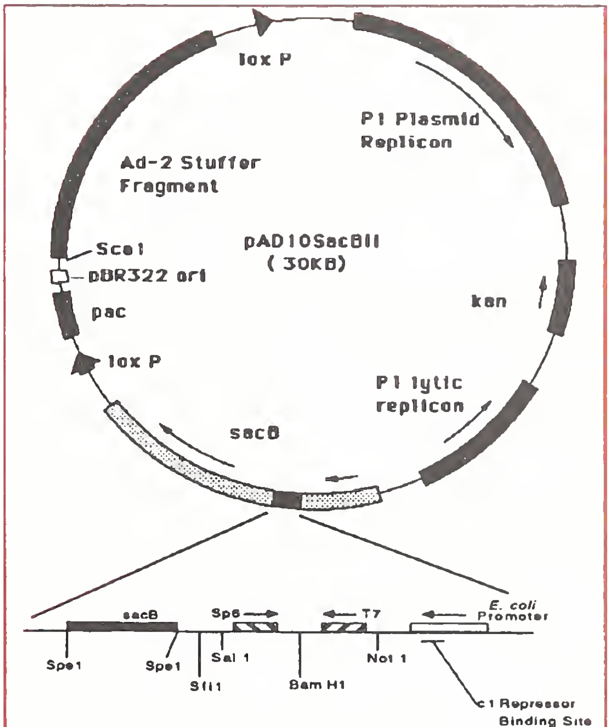


Fig. 1
Plasmid map of pAd10-SacB II and cloning region

To overcome some of the problems associated with using cosmid or YAC systems, a novel method for cloning and packaging DNA fragments using a Bacteriophage P1 system has been developed[3] that offers the ability to clone large genomic DNA fragments of between 70-95 kb in size with efficiencies approaching those of cosmids. In addition, the P1 DNA Packaging System uses host *E. coli* strains and *in vitro* packaging extracts obtained from strains that are deficient in restricition and recombination abilities. These prevent the degradation and recombination of methylated genomic DNA.[4]

Principle of the Method

Using a strategy analogous to Phage λ packaging,[1] partially digested and size selected genomic DNA between 70 and 95 kb is ligated onto linearized plasmid vector arms. The SacBII vector used contains a Phage P1 pac-cleavage site and two Phage P1 loxP recombination sites in addition to replication origins and an antibiotic resistance gene (Fig. 1) The recombinant vector is cleaved at the pac-cleavage site in a "Pacase" extract and the resulting DNA is then inserted into an empty P1 Phage head using a second extract containing phage packaging proteins. The attachment of P1 phage tails to the heads results in the formation of infectious recombinant phage particles that are then used to infect a restriction minus *Escherichia coli* host strain containing an expressed cre gene. ▶

After injection into the host strain, the recombinant DNA is circularized between the two Phage P1 loxP sites by Cre recombinase. DNA which does not circularize is degraded by host nucleases. The circular DNA molecule now replicates and is maintained stably at one copy per host cell by the P1 Plasmid replicon.

Prior to alkaline lysis plasmid isolation,[5] the recombinant plasmid copy number is increased more than 25-fold by isopropyl B-D-thiogalactopyranoside (IPTG) induction of the lac promoter controlled high-copy P1 lytic replicon. A schematic detailing the P1 DNA packaging strategy is shown in Figure 2.

Phage P1 uses a headful packaging strategy. Once the phage head is filled with DNA (about 110-115 kb), a "headful cut" occurs, cleaving any remaining DNA away from the head before it is packaged. This packaging mode suggests that if the insert DNA is too large (>95 kb), it will be packaged but not recovered in bacteria because the "headful" packaging process will terminate before the distal loxP is incorporated into the phage head. It also suggests that DNA too small to generate a headful when inserted into the vector should not be packaged into phage particles.

However, it has been observed that DNA less than a P1 "headful" is packaged *in vitro* in viable phage particles with an efficiency of about 10-15 percent that of "headful" DNA (Pierce & Sternberg, manuscript in preparation). While it is not completely clear why this should occur, it points out the need to size-select the insert DNA on sucrose gradient before attempting to ligate it to vector in order to maximize the recovery of large inserts. In particular, size fractions that maximize DNA fragments in the 70-95 kb range, and minimize the presence of shorter fragments, should be used. The presence of small fragments is

Developments

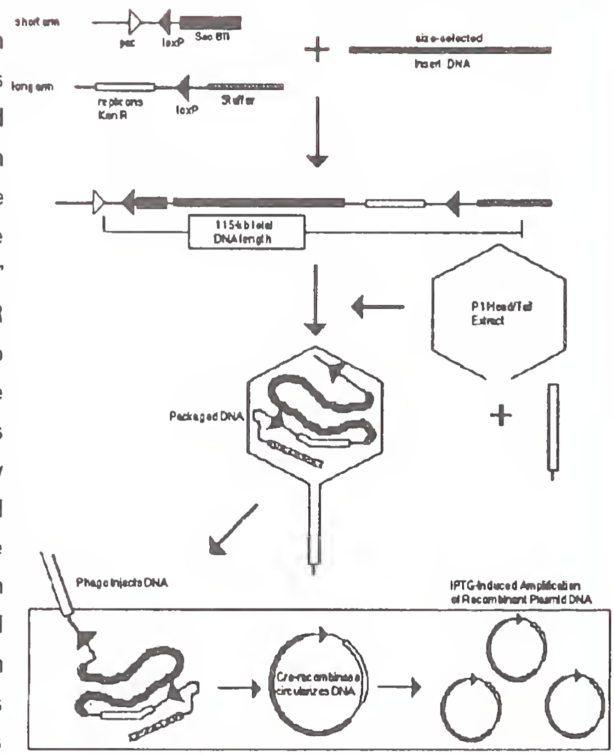
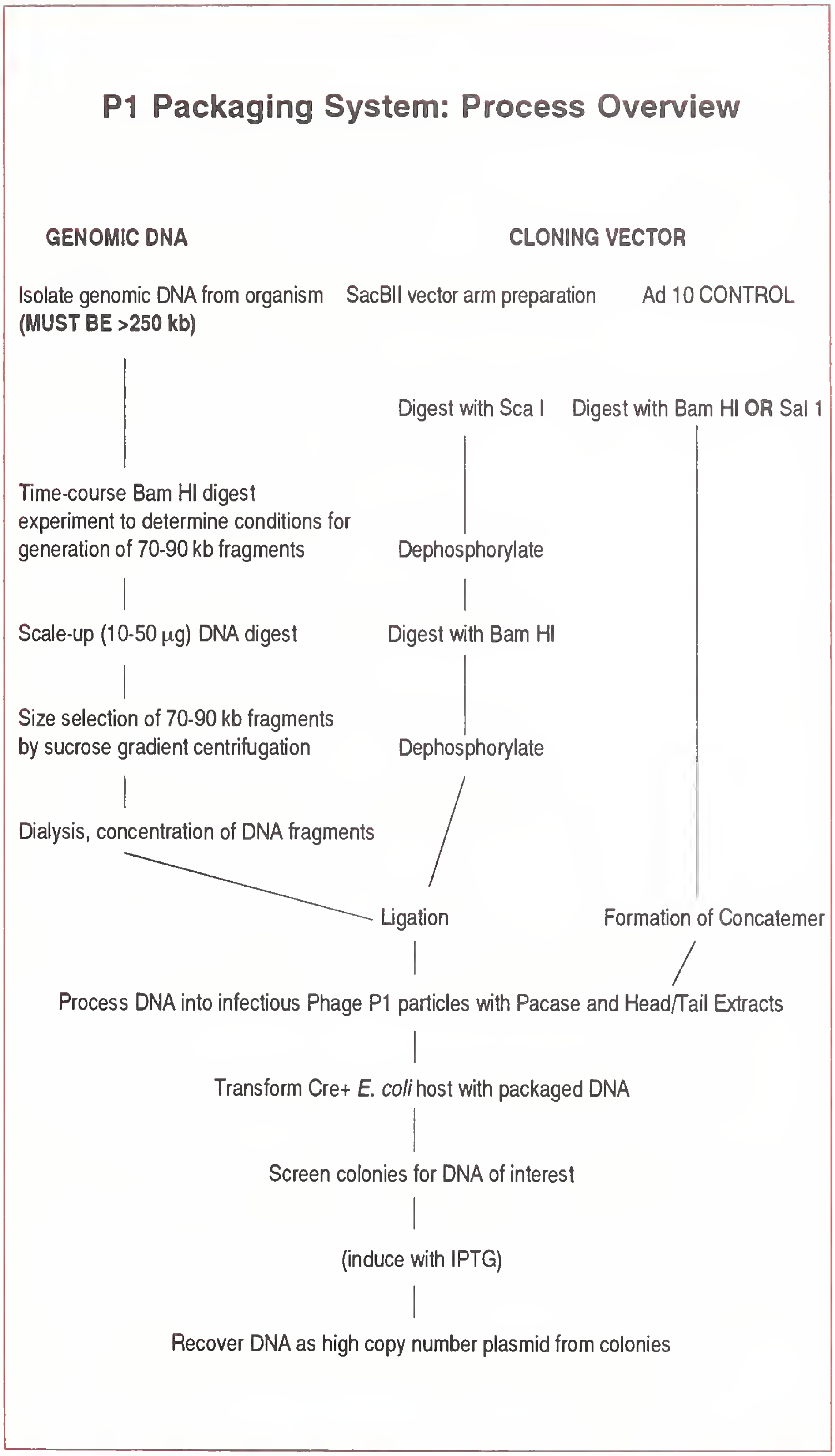


Figure 2. The P1 DNA Packaging Strategy. The SacBII vector is digested with restriction endonucleases ScaI and BamHI and the ends dephosphorylated. This generates two vector "arms," one consisting of the "short" ScaI to BamHI fragment and the other the "long" ScaI to BamHI fragment. Genomic DNA is partially digested with BamHI or other compatible end restriction endonucleases and size-selected on a sucrose gradient. Fragments between 70 kb and 95 kb in length are isolated and ligated to the vector arms, generating a series of linear molecules. If ligation occurs between two "short" arms, the resulting molecule will neither contain the origins of replication nor the kanR gene, and will be nonviable. If both arms are "long," there will be no pac site, and no packaging into the phage heads will occur. The only viable recombinant will be one consisting of the insert sequence flanked by both a short and long arm. Phage P1 uses a headful packaging strategy and can accommodate a total DNA length of approximately 110-115 kb. Any inserts longer than 95-100 kb will result in truncation of the packaged DNA before the distal loxP site is inserted, and the molecule will be unable to circularize upon injection into the host. Once injected into the cre+ host cell, the cre protein circularizes the injected DNA at the loxP sites, and DNA now replicates using the plasmid origin of replication. Propagation of cells on sucrose containing media only permits growth of colonies with genomic DNA inserts (positive selection). Plasmid copy number is increased by induction with IPTG. The recombinant DNAs are then isolated as plasmids using traditional methods.



undesirable also because it increases the possibility that they will be ligated together and then recovered in one clone. This would significantly complicate subsequent screening and gene localization processes.

Finally, as the second stage packaging extract contains about 5-10 percent small heads (headful size 47 kb), small DNA fragments can be recovered by a headful packaging process in phage particles containing these heads.

An 11 kb "stuffer" region of Adenovirus type 2 DNA has been engineered into the pAd10-SacBII cloning vector. It is designed to provide a segment of DNA in which the "headful" cut can be made.

Using the P1 DNA Packaging System, genomic DNA from 70-95 kb can be readily cloned and manipulated. The major advantages of the P1 DNA packaging method over other genomic cloning methods are: 1) the ability to clone inserts two to three times the size of those used with cosmids and lambda vectors, 2) no rearrangement or deletion of methylated DNA occurs because of the use of restriction-minus host strains, and 3) recombinant DNA is easily recovered as plasmids for further screening and minipulations.

P1 Cloning Vectors

One of the key components in the construction of a P1 DNA library is the genetically engineered P1 plasmid vector. The original vector was pNS582tet14Ad10 (see Fig. 3). A new improved, more versatile version of the Ad10 vector has been constructed. This vector is designated pAd10-SacBII (see Table 1) and will be available soon. ◆

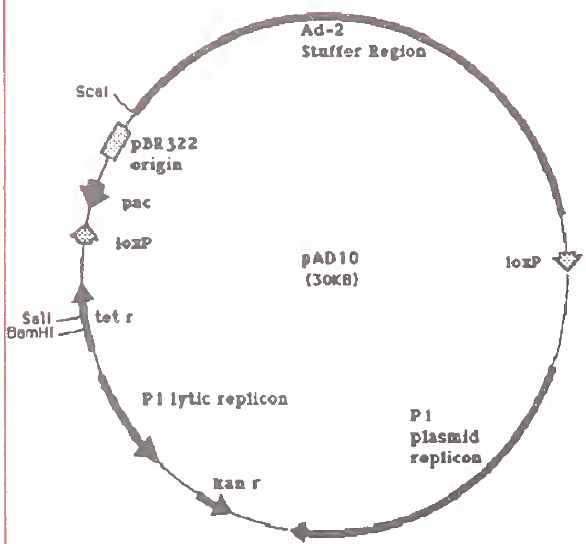


Fig. 3 Plasmid map of pNS582tet14Ad10

Table 1. Features of SacBII and Ad10 Plasmid Vectors

	SacBII	Ad10
Cloning	*Bam*HI*	*Bam*HI / *Sal* I
Screening Method	positive Kan/sucrose	antibiotic Kan/Tet
Rare 8-base R.E sites	*Sfi* I/*Not* I	none
RNA Analysis	Sp6/T7	none

**Sal* I can also be used to clone inserts, but this prevents use of Sp6 and T7 promoters to prepare RNA-probes from cloned insert.

REFERENCES:

1. Murray, N.E. (1983) in *Lambda II*, eds. Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.) pp.395-432.
2. Burke, D.T., Carle, G.F. and Olson, M.V. (1987) *Science* 236: 806-812.
3. Sternberg, N.L. (1990). *Proc. Natl. Acad. Sci.* USA 87:103-107.
4. Blumenthal, R. (1989). *Focus* 2:41-46.
5. Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* 7:1513-1523.
6. Sternberg, N.L. (1990) *Gen. Analysis Tech. Applications* 7:126-132.
7. Sternberg, N., Reuther, J. and deRiel, K. (1990) *The New Biologist* 2:151-162.
8. Pierce, J.C. and Sternberg, N.L. (In Press) *Methods in Enzymology*.
9. Pierce, J.C., Sauer, B.S. and Sternberg, N.L. (In Press) AMAS.

Du PONT'S P1 DNA Packaging System

ORDERING INFORMATION

Call 1-800-551-2121 1,1 at the prompt

NEP-113

NEP-113 contains all the components found in both NEP-113C and NEP-113P, listed below. **(NOTE: NEP-113 does not include "SacBII" positive selection cloning vector.)**

NEP-113C

The P1 DNA Cloning System contains the components needed for ligating size-fractionated genomic DNA into the P1 "SacBII" cloning vector (ordered as **NEP-113V**). The kit includes: P1 control "Ad-10" vector, T4 DNA ligase and ligation buffer, Calf Intestinal Alkaline Phosphatase, and high molecular weight DNA markers (used for selecting DNA of optimal length for cloning and packaging).

NEP-113P

The P1 DNA Packaging System consists of all components needed for five packaging reactions of genomic DNA to P1 vector clones, including the phage P1 Head/Tail extract, the Pacase extract for cleaving the recombinant DNA and guiding it into phage heads, the recombination and restriction-minus bacterial plating host strains, and all necessary buffers.

NEP-113V

New P1 cloning vector pNS582tet14-Ad10-SacBII, the positive selection cloning vector used for genomic DNA library generation. **Must be ordered separately; not included with NEP-113 or NEP-113C. Available soon.**

On the Horizon

Calendar of Upcoming Genome Events

Meetings - 1992

January 25-February 1: **Molecular Mechanisms In DNA Replication & Recombination,** Taos, NM. Contact: Keystone Symposia, Drawer 1630, Dept. S, Silverthorne, CO 80498. Telephone: (303) 262-1230, FAX: (303) 262-1525.

March 15-18: **30th Annual Meeting of the American Cytogenetics Conference,** Virginia Beach, VA. Contact: A. Brothman, Telephone: (804) 446-5670, FAX: (804) 624-2255.

March 19-22: **Maize Genetics Conference,** Asilomar State Park, Pacific Grove, CA. Contact: S. Hake, Telephone: (510) 559-5922, Fax: (510) 559-5678.

April 27-28: **Annual Biotechnology Patent Conference,** Washington, DC. Contact: ATCC Workshop Manager, Telephone: (301) 231-5566, FAX: (301) 770-1805.

May 6-8: **6th Annual Seminar on Analytical Biotechnology,** Cambridge, MA. Contact: Barr Enterprises, P.O. Box 279, Walkerville, MD. Telephone: (301) 898-3772, FAX: (301) 898-5596.

May 17-20: **Association of Biotechnology Companies Sixth International Biotechnology Meeting,** San Diego, CA. Contact: ABC Headquarters for Registration and Membership Information, ABC, 1666 Connecticut Ave, NW, Suite 330, Washington, DC. Telephone: (202) 234-3330, FAX: (202) 234-3565.

May 28-30: **Penn State Symposium in Plant Physiology: Biosynthesis and Molecular Regulation of Amino Acids in Plants,** University Park, PA. Contact: Dr. Jack C. Shannon, 102 Tyson Bldg, University Park, PA 16802. Telephone: (814) 863-2192, FAX: (814) 863-6139.

June 22-24: **Annual Meeting of the Electrophoresis Society,** Research Triangle Park, NC. Contact: Barr Enterprises, P.O Box 279, Walkerville, MD. Telephone: (301) 898-3772, FAX: (301) 898-5596.

August 4-8: **11th International Chromosome Conference,** Edinburgh, UK. Contact: Dr. Ann Chandley, MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, United Kingdom.

August 16-21: **9th International Biotechnology Congress,** Crystal City, VA. Contact: Congress Office, American Chemical Society, 1155 Sixteenth Street, N.W. Room 205, Washington, DC 20036.

August 23-28: **8th International Symposium on Yeasts,** Atlanta, GA. Contact: Dr. S.A. Meyer, Dept. of Biology, Georgia State University, P.O. Box 4010, Atlanta, GA 30302.

Workshops and Courses - 1992

February 3-21: **ICGEB-UNIDO Workshop 1992: Gene Isolation and Analysis for Crop Improvement,** New Delhi, India. Contact: Mr. G. Chatterjee, Administrative Officer, ICGEB, NII Campus, Shahid Jeet Singh Marg, New Delhi 110 067, India. Telephone: 91-11-6887353, FAX: 91-11-6862317.

March 2-17: **Carolina Workshop on Yeast Molecular Genetics,** Chapel Hill, NC. (application deadline: Feb 1). Contact: W. Litaker, (919) 966-1730, FAX: (919) 966-6821.

April 27-29: **3rd European Workshop on Cytogenetics and Molecular Genetics,** Porto, Portugal. Contact: Dr. S. Catedo, Dept. of Medical Genetics, Medical Faculty of Porto, Portugal. Telephone: 351-2-497833, FAX: 351-2-4103940.

June 1992: **Genomic Information: Ethical Implications,** Seattle, WA. Contact: B. Brownfield. Telephone: (206) 543-5447, FAX: (206) 685-7515.

Future Events

April 21-25 1993: **Molecular Genetics of Plant-Microbe Interactions (Workshop and Symposia),** East Brunswick, NJ. Contact: Rutgers, The State University of New Jersey, Registration Desk, Office of Continuing Professional Education, Cook College, P.O. Box 231, New Brunswick, NJ 08903. Telephone: (908) 932-9271, FAX: (908) 932-8726.

May 8-13 1994: **HPLC'94, Eighteenth International Symposium on High Performance Liquid Chromatography,** Minneapolis, MN. Contact: Barr Enterprises, P.O. Box 279, Walkerville, MD. Telephone: (301) 898-3772, FAX: (301) 898-5596.

June 16-21 1996: **HPLC'96, Twentieth International Symposium on High Performance Liquid Chromatography,** San Francisco, CA. Contact: Barr Enterprises, P.O. Box 279, Walkerville, MD. Telephone: (301) 898-3772, FAX: (301) 898-5596.

1992 MAIZE GENETICS MEETING

March 19–22, 1992

FEATURED SPEAKERS

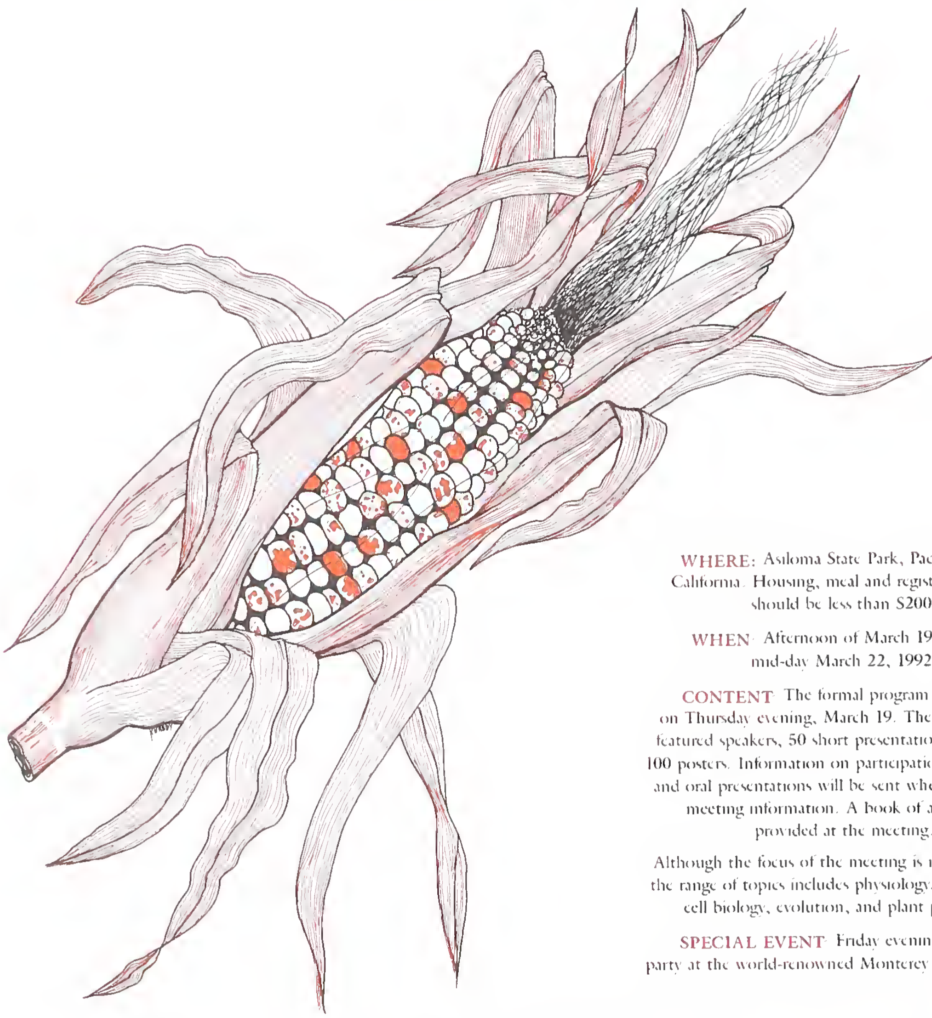
Don McCarty, University of Florida
The role of Viviparous-1 in regulating seed maturation

Peter Starlinger, Institut für Genetik, Köln
Open questions on the plant transposable element Ac

Vicki Chandler, University of Oregon
Molecular and genetic analysis of the anthocyanin pathway

Sheila McCormick, USDA Plant Gene Expression Center
Functional analysis of gene expression in pollen

Tony Pryor, C.S.I.R.O., Canberra
Is evolution hyperactive at rust resistance loci?



WHERE: Asiloma State Park, Pacific Grove, California. Housing, meal and registration costs should be less than $200

WHEN: Afternoon of March 19 through mid-day March 22, 1992.

CONTENT: The formal program will begin on Thursday evening, March 19. There will be five featured speakers, 50 short presentations, and about 100 posters. Information on participation in the poster and oral presentations will be sent when you request meeting information. A book of abstracts is provided at the meeting.

Although the focus of the meeting is maize genetics, the range of topics includes physiology, development, cell biology, evolution, and plant pathology.

SPECIAL EVENT: Friday evening tour and party at the world-renowned Monterey Bay Aquarium.

For more information contact:
Ed Coe, USDA-ARS, Department of Agronomy, Curtis Hall, University of Missouri, Columbia, MO 65211

Transportation: The Park is 7 miles from the Monterey Airport, served by local transport. For those arriving in the San Francisco Bay Area, a rental car will be required. The Park is approximately 1.5 hours from the San Jose Airport and 2.5 hours from both the San Francisco and Oakland Airports.

Columbus Ends Global Isolation

Dr. Susan McCarthy, Coordinator
Plant Genome Data and Information Center
National Agricultural Library, USDA
Beltsville, MD

When Christopher Columbus set sail in 1492 he linked two old worlds. Ten thousand years had separated them. Columbus began the process of encounter and exchange that continues to unfold today. The U.S. Smithsonian Institution has created an ambitious exhibit, "Seeds of Change," that focuses on the biological and cultural impacts initiated 500 years ago.

"Seeds of Change" presents an unusual perspective of that event. Rather than focusing on deliberate human actions as instigators of world events, the exhibition traces five biological transfers and their effects on both the old and new worlds. The exhibition approaches history not as a linear process, but as interrelationships. It presents a new way of looking at world history by examining five "seeds"--corn, potato, sugar, horse, and disease--and their key roles in shaping human history.

14 92
19 92

SEEDS of CHANGE

QUINCENTENARY.

NATIONAL MUSEUM OF NATURAL HISTORY
SMITHSONIAN INSTITUTION

Constitution Avenue at 10th Street NW
Washington DC 20560

Panel versions of "Seeds of Change," co-sponsored by the American Library Association and the Smithsonian Traveling Exhibition Service (SITES), will be exhibited at public libraries in all 50 States and 3 U.S. territories. SITES is also circulating a copy to museums throughout the country.

The National Museum of Natural History, Office of Education, will be presenting lectures, films, festivals, demonstrations, and tours from October 26, 1991, through April 1993. For more information, please call (202) 357-2747. ◆

Test your knowledge ▶

NAL Kicks Off the Traveling Show

The traveling "Seeds of Change" panel version began its national tour at the National Agricultural Library (NAL). The exhibit opened December 9, 1991, in conjunction with the American Library Association, Exhibit Workshop. The traveling exhibit will reach 60 libraries over the next 2 years.

"Seeds of Change" Quiz

1 Today corn and potatoes supply how much of the global calories?

- a) 5 percent
- b) 10 percent
- c) 20 percent
- d) 40 percent
- e) 80 percent

2 Conquest of the New World caused a loss in native populations. Between 1492 and 1900 how great was the loss?

- a) 3 percent
- b) 9 percent
- c) 27 percent
- d) 54 percent
- e) 90 percent

3 While New World populations were on the decline Old World populations were rapidly increasing. In the same time frame how large was the European population increase?

- a) 20 percent
- b) 80 percent
- c) 120 percent
- d) 240 percent
- e) 444 percent

4 New World herbal remedies included the following drugs: quinine; cocaine; and cortisone. Which New World botanical provided cortisone?

- a) yams
- b) cassava
- c) maize
- d) peanuts
- e) squash

5 The New World has incorporated many waves of immigration. The African continent is the source of the single largest forced migration. Between 1492 and 1888 how many Africans were brought to this hemisphere?

- a) 1 million
- b) 2 million
- c) 5 million
- d) 10 million
- e) 20 million

6 Which labor-intensive crop introduced slavery to the New World?

- a) cotton
- b) sugar cane
- c) tobacco
- d) chocolate
- e) coffee

7 Which of the following crops was introduced from the Old World?

- a) pineapple
- b) rubber
- c) sugar cane
- d) chili peppers
- e) vanilla

8 How long, in days, was the first voyage of Columbus?

- a) 33
- b) 54
- c) 67
- d) 89
- e) 121

9 How many voyages to the New World did Columbus make?

- a) one
- b) two
- c) three
- d) four
- e) five

10 The New World has many resources. What has become the most enduring and important resource exchanged with the Old World?

- a) golden artifacts
- b) Colombian emeralds
- c) the Fountain of Youth
- d) botanical wealth
- e) land

QUIZ ANSWERS

1. d; 2. e; 3. e; 4. a; 5. d; 6. b; 7. c; 8. a; 9. c; 10. d

Theft of Threatened Plants Hinders Recovery Effort

Chris Holmes and Phil Villa-Lobos
USDA, Office of Press and Media Relations
News Division
Washington, DC

Washington, Aug. 5--A theft of rare plants from a national forest in Illinois may significantly hinder efforts to save the threatened Mead's milkweed, according to the U.S. Department of Agriculture's Forest Service.

Forest Service Chief F. Dale Robertson said law enforcement officials from the Forest Service and U.S. Fish and Wildlife Service are investigating the plant theft. The Shawnee National Forest has offered a $5,000 reward for information leading to the arrest and conviction of the plant thieves.

Robertson said an entire population of Mead's milkweed was stolen from the Shawnee National Forest in mid-June, including both wild plants and young plants recently planted through a cooperative reintroduction effort. The plant is listed by the Federal Government as threatened under the Endangered Species Act.

Mead's milkweed is an important remnant of the tallgrass prairie which was once prevalent in much of the Midwest, though much of its habitat has been converted to other uses. The Shawnee National Forest site was one of only six places where the plant was known to exist in the Eastern United States, and is the "premier site for reintroducing the plant in its native habitat," Robertson said.

Mead's milkweed has been the focus of a national cooperative recovery effort involving the Forest Service; the U.S. Fish and Wildlife Service; the Morton Arboretum in Lisle, IL; and the Illinois Department of Conservation.

Robertson said that botanists fear the theft of the plant may have compromised the recovery effort for the species. "This is a major setback for our cooperative efforts to reestablish the species east of the Mississippi River," Robertson said. "We're going to have to redouble our efforts if we're going to be successful." He said additional Mead's milkweed plants grown at the Morton Arboretum will be replanted at the damaged site, and other possible sites for the milkweed will be restored and planted as well. ◆

Mead's milkweed: Efforts continue to reintroduce this threatened plant in the Shawnee National Forest.

USDA Photo

Introducing Dr. Douglas Bigwood

Dr. Douglas Bigwood is the Database Manager for NAL's Plant Genome Data and Information Center. Dr. Bigwood is primarily responsible for directing the implementation of the plant genome database system at NAL. The database is an initiative of USDA's Plant Genome Research Program. Dr. Bigwood coordinates database activities of the Center and USDA's Plant Genome Research Program in the participating agencies. He also coordinates this database work with that of other genome research programs outside USDA. Dr. Bigwood was first employed with the Center in February 1991 through a contract with the University of Maryland.

Prior to coming to NAL, Dr. Bigwood was self-employed as a consultant in expert systems and neural network development. From 1989 to 1991, he worked as a consultant for the Network Technology Division at COMSAT Laboratories. His accomplishments there included developing a Network Management Expert System.

Dr. Bigwood has worked previously with USDA's Agricultural Research Service. From 1985 to 1989, he was employed as a contract scientist with the Systems Research Laboratory. Responsibilities there included the research and development of artificial intelligence technology.

From 1984 to 1985, Dr. Bigwood served as a senior systems analyst for the ARS Technology Transfer Office. While in this position, he developed the Technology Transfer Automated Retrieval System (TEKTRAN), which allows rapid access to new research results submitted for publication.

Dr. Bigwood initially worked with ARS as a microcomputer specialist in the Communications and Data Services Division, There he evaluated new microcomputer hardware and software, and designed and developed computer programs.

Dr. Bigwood has had experience with a large number of operating systems and programming languages, and a variety of hardware and software. He has authored or coauthored numerous articles in scientific journals and proceedings on expert systems, artificial intelligence, and database management systems. In Addition, Dr. Bigwood has been an invited speaker at several scientific and professional conferences. He is a member of various professional organizations, including the American Association for Artificial Intelligence and the International Neural Network Society.

A native of Massachusetts, Dr. Bigwood earned his B.S. degree in botany from the University of New Hampshire. He holds a Ph.D. in plant ecology, with minors in genetics and statistics, from the University of Maryland. ◆

Probe

ISSN: 1057-2600

The official quarterly publication of the USDA Plant Genome Research Program. This newsletter is aimed at facilitating interaction throughout the plant genome mapping community and beyond.

Probe is a publication of the Plant Genome Data and Information Center, National Agricultural Library.

Managing Editor
Susan McCarthy, Ph.D.

Editor
Carolyn Bigwood

Assistant Editor
Deborah Richardson, M.S.

Production Manager
Terrance Henrichs

Layout and Design
Barbara Binder
Terrance Henrichs

Special Thanks to:
Nancy Domchick
Barbara Buchanan
Kathleen Hayes

Articles, announcements, and suggestions are welcome.

Correspondence Address
Susan McCarthy
NAL, Room 1402
10301 Baltimore Blvd
Beltsville, MD 20705
Phone:(301)504-6875
FAX: (301)504-7098

USDA Program Office
Dr. Jerome Miksche
USDA/ARS/NPS/PNRS
Room 331C, Bldg 005
BARC-WEST
Beltsville, MD 20705
Phone:(301)504-6029
FAX: (301)504-6231

NAL
National Agricultural Library

Probe

fold

Place Stamp Here

Plant Genome Data and Information Center
USDA - National Agricultural Library
10301 Baltimore Blvd., Room 1402
Beltsville, Maryland 20705-2351

fold

Staple **Staple** **Staple**

If you would like to continue receiving this publication, please return this page. If we do not receive a response within three issues your name will be removed from the mailing list.

- ☐ Keep my name on the mailing list.
- ☐ Correct name on the label.
- ☐ Correct address on the label.
- ☐ Replace name on the label with my name.
- ☐ Add my name to the mailing list, but do not delete the name on the label.
- ☐ Remove my name from the mailing list.

Name:

Address:

City: State: Zip:

Telephone: ( ) - FAX: ( ) -

Photocopies of this slip are acceptable. Please submit entire page.

Plant Genome Data
and Information Center
USDA - NAL
10301 Baltimore Blvd.
Beltsville, Maryland 20705-2351

ISSN: 1057-2600

BULK RATE
POSTAGE AND FEES PAID
USDA
PERMIT NO. G95